



Statistical Analysis of Weather Forecast Data: Wunderground Forecasts vs. Statistical Models

Sijia Huo, Yunxin Zhang, Yantong Zheng, Yayu Zhou; Faculty Advisor: A.J. Hildebrand (University of Illinois at Urbana-Champaign)

Undergraduate Poster Session at the Joint Mathematics Meetings, Atlanta, GA, January 6, 2017



PREVIOUS WORK

- **Accuweather Long-Range Forecast Accuracy (Rosenberg, 2013)** - Forecasts collected from Accuweather over a 6-month period found to be worse than climatology past 10 days.
- **Accuracy of Official Weather Forecasts in Melbourne (Stern, 2008)** - Melbourne providing an environment challenging to forecasters due to relatively unpredictable climate. Analysis conducted on forecast data from 1960 to 2008.

DESCRIPTION OF DATA

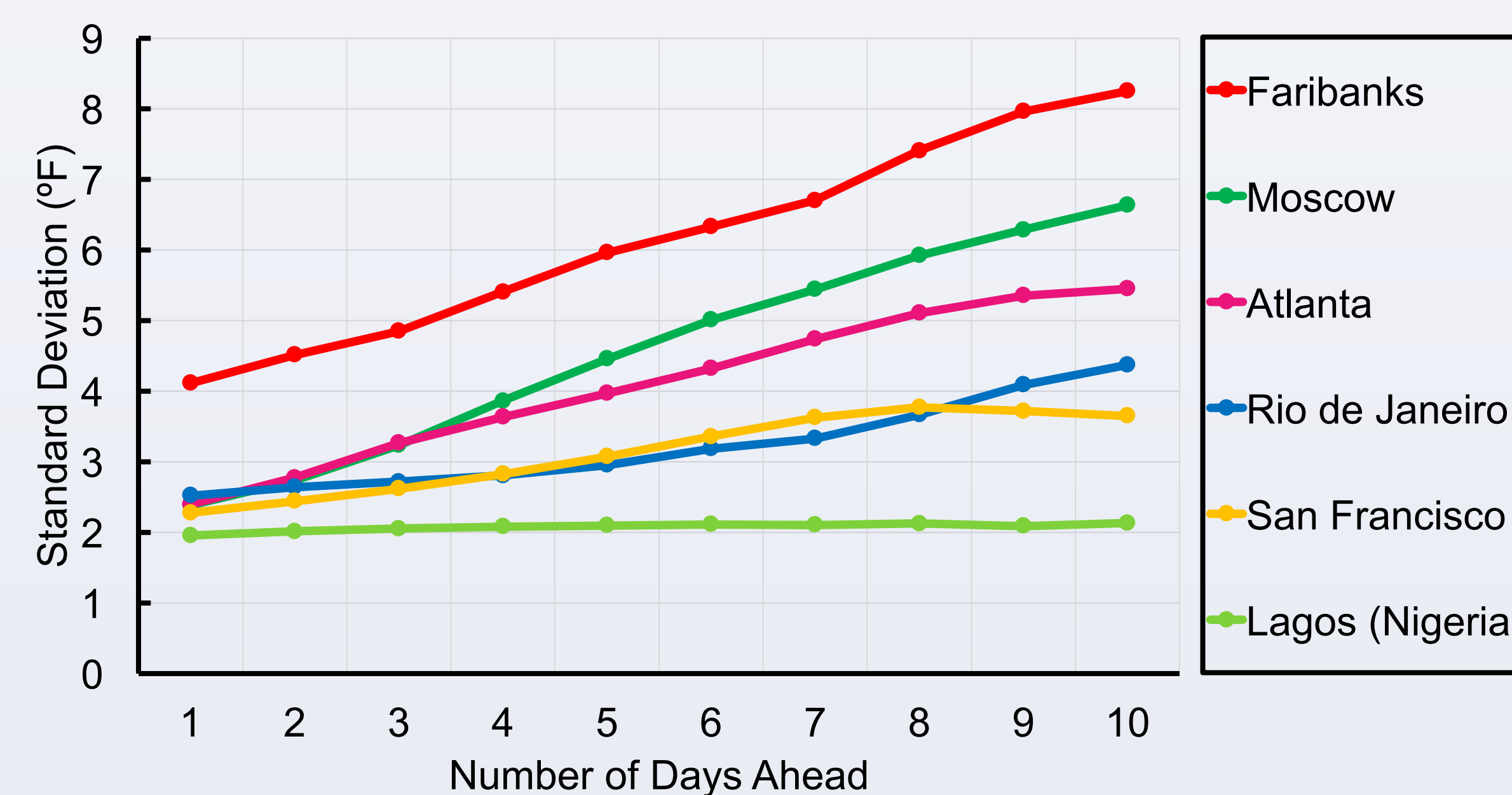
- **Historical Data:**
 - Historical data downloaded from the Utah State Climate Center from <https://climate.usurf.usu.edu>
 - Daily maximal and minimal temperatures for up to 100 years from 4,625 locations in the U.S.
 - Over 100,000,000 data points altogether
- **Wunderground Forecast Data:**
 - Forecast data collected each hour with API key from Wunderground.com using a Python script
 - Hourly forecasts up to 240 hours out collected from May 2016 to December 2016 for 33 selected cities around the world, including Atlanta and San Francisco
 - Over 1,200,000 predictions collected for each city

METHODOLOGY

- **Data Validation:** missing data that appear at the beginning and the end of the data set were deleted; missing data in the middle were replaced by the value of the one-day ahead data (if not missing) or that of the one-day behind data
- **Metric:** standard deviation used to measure the accuracy of different methods
- **Persistence Model:** a model that assumes that the forecast temperature is the same as the current temperature
- **Climatology Model:** a model that uses the historical average temperature for each day of the year as the temperature prediction for that day
- **Multiple Linear Regression (MLR) Model:** a linear regression model using temperatures of the previous 3 days as predictors
- **Time Series Model:** an ARIMA model with parameters (3, 0, 3), which assumes the data to be stationary

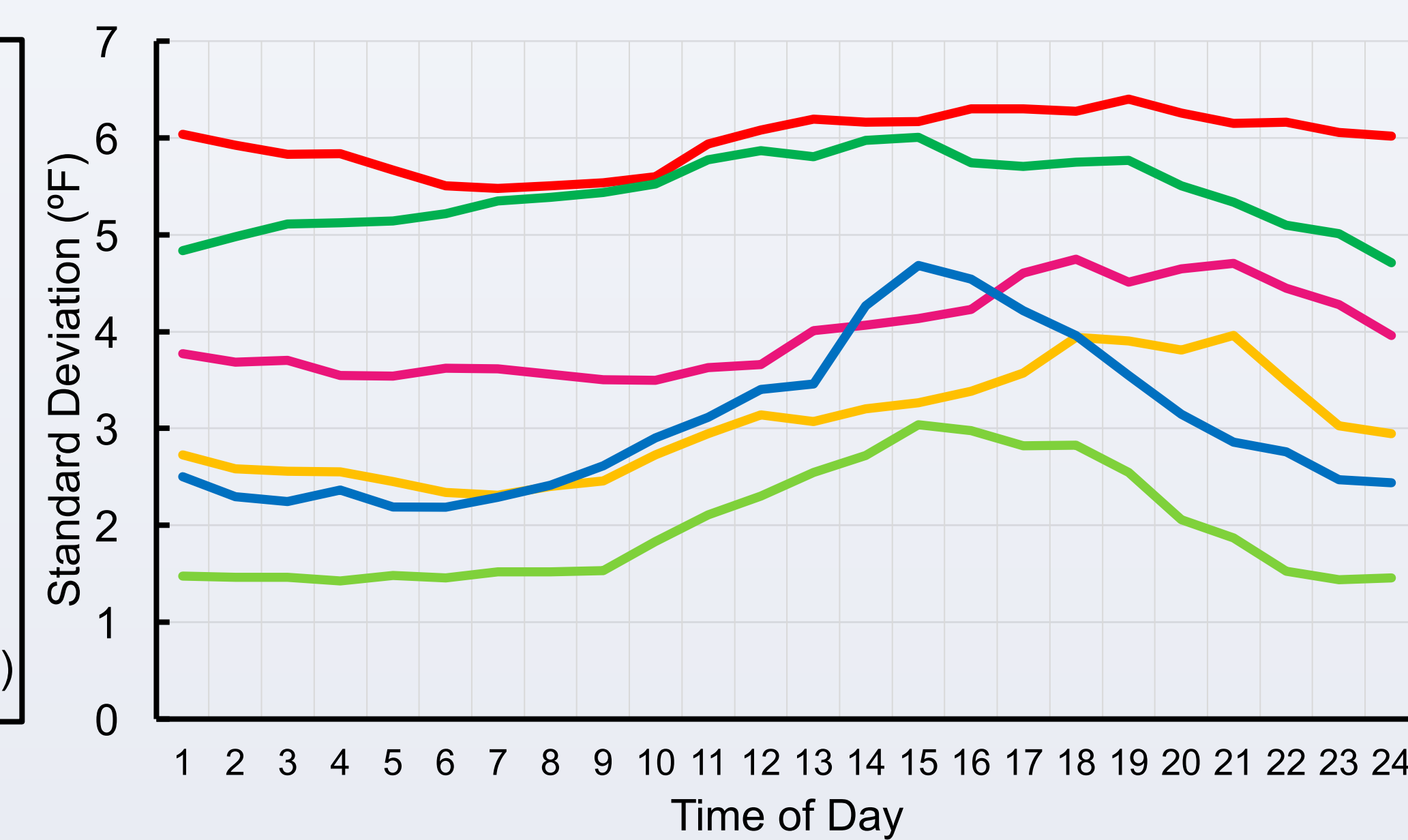
ACCURACY OF WUNDERGROUND FORECASTS

AVERAGE FORECAST ERROR BY DAYS AHEAD



- Average error (measured as standard deviation between actual maximal temperatures and forecasts) for Lagos is consistently less than other locations

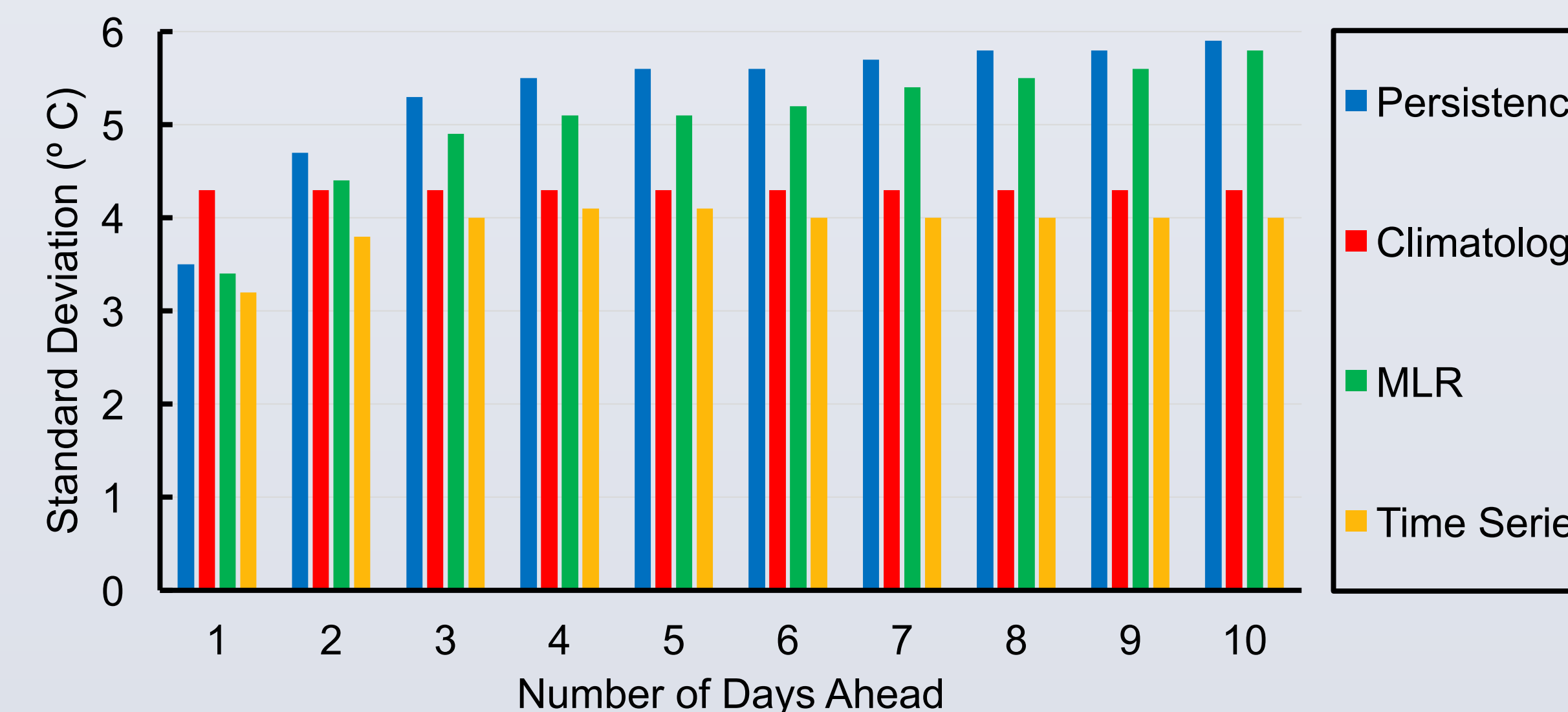
AVERAGE TEN DAY FORECAST ERROR BY TIME OF DAY



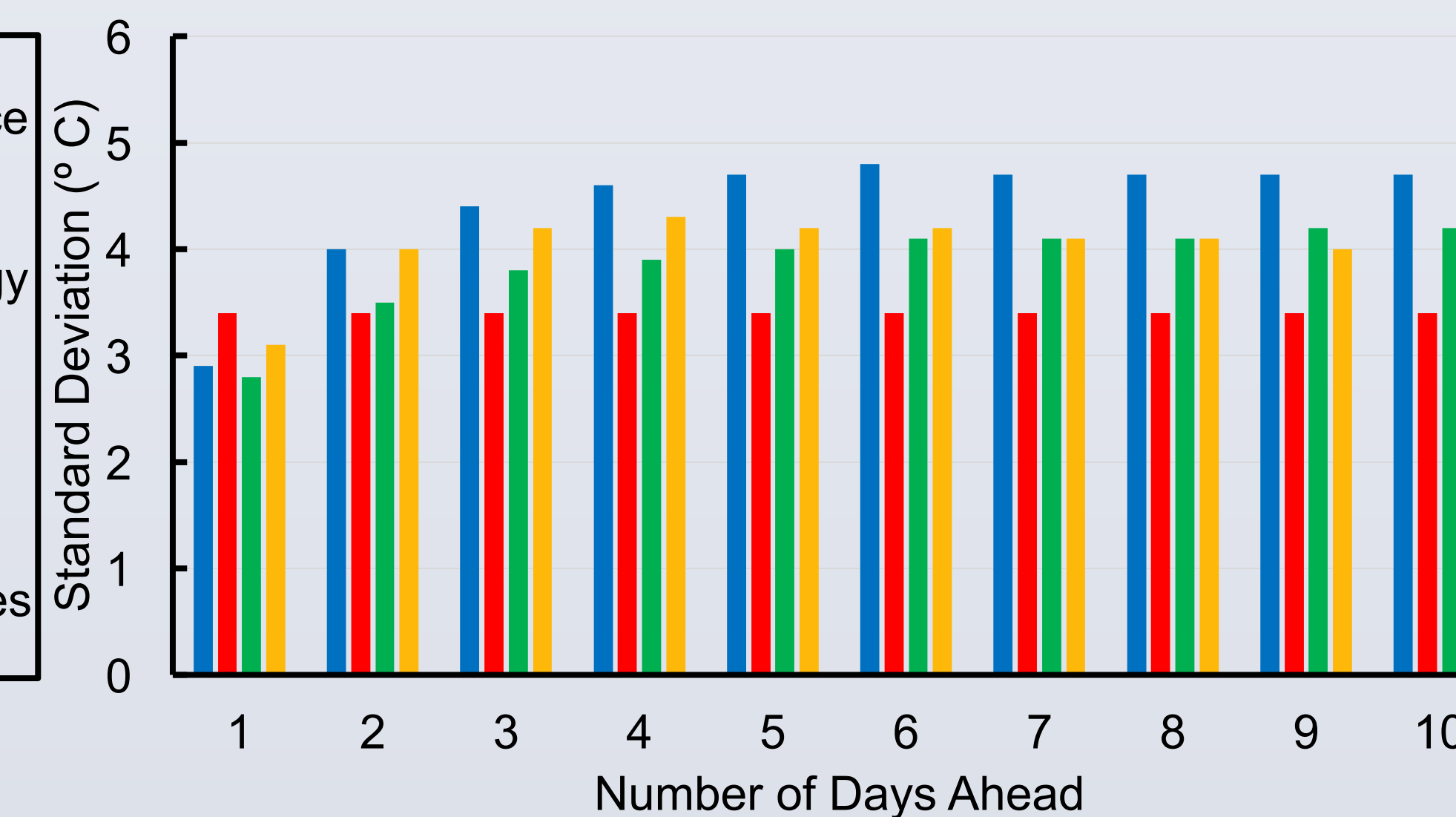
- Hourly average errors were the highest in the afternoon (1pm - 6pm) and the lowest in the morning (4am - 9am).

ACCURACY OF STATISTICAL MODELS

ATLANTA

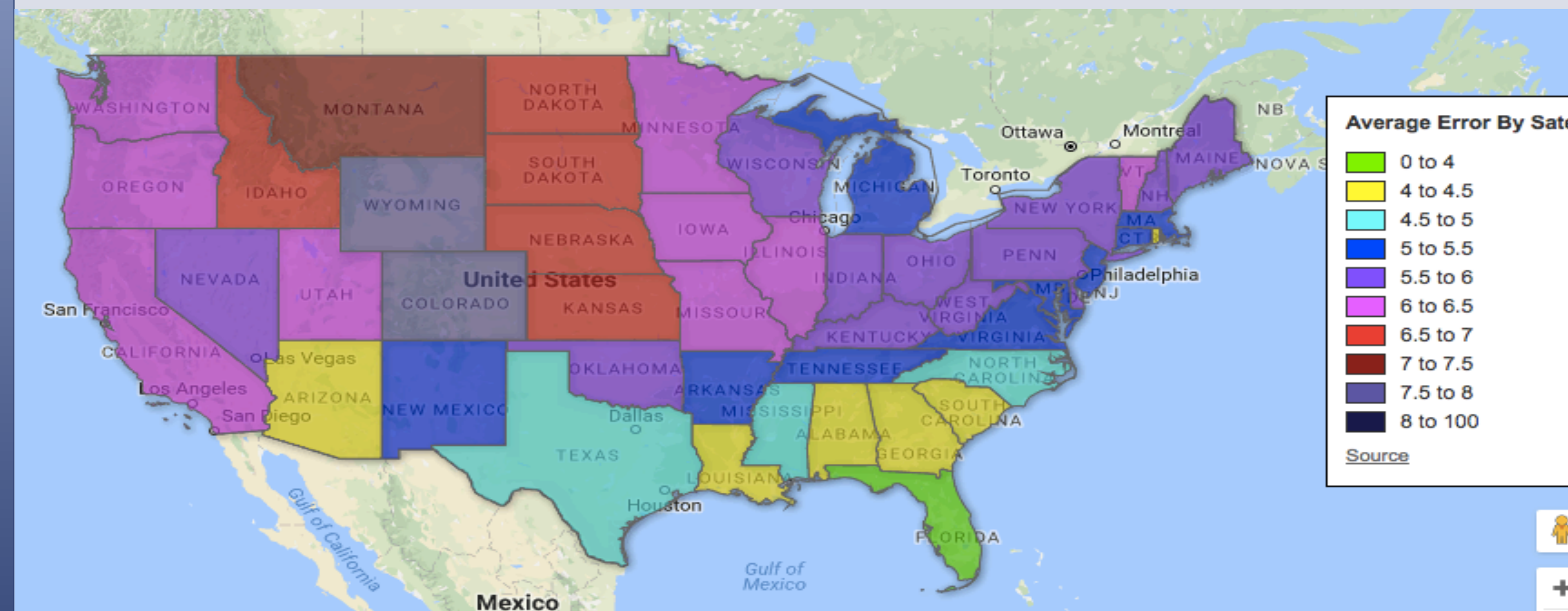


SAN FRANCISCO



- **Accuracies for statistical models:** All models performed better in San Francisco than in Atlanta; the Persistence model performed better than the Climatology model on day 1 and worse on days 2-10.

ACCURACY OF CLIMATOLOGY MODEL BY STATE



- The climatology error is smallest for Hawaii (std. dev. 1.8 °C) and Florida (3.4 °C), and largest for Montana (7.2 °C), Wyoming (7.7 °C) and Colorado (7.2 °C).

SUMMARY OF OBSERVATIONS

- **Wunderground forecasts vs. statistical models:** The Wunderground forecasts were on average significantly more accurate than the statistical model predictions up to seven days ahead (at the 1% significance level).
- **Accuracy of statistical models:** An optimal linear combination of the Persistence and Climatology models (Murphy, 1992) performed better than the Persistence and Climatology models individually, but was not significantly different from the MLR and Time Series models (at the 1% level).
- **Accuracy of long-term forecasts:** In most locations, the average forecast errors increased at a linear rate from one day ahead to ten days ahead. In some locations (for example, Lagos and San Francisco), the ten-day forecasts were as accurate as the one-day forecasts.
- **Accuracy at different locations:** Within the U.S., all models perform better in southern states than in northern states. Among international cities, the accuracy was best at locations near the equator (for example, Lagos).

FUTURE WORK

- Extend the analysis to other variables, for example, precipitation, cloud cover, etc.
- Extend the analysis to other forecast sources, for example, the National Weather Service and AccuWeather.com.
- Investigate different measures of accuracy, for example, maximal forecast errors, mean absolute errors, and confidence intervals.
- Scale up the analysis to more locations.

REFERENCES

- **J.D. Eggleston and S. Dubner**, *How Valid are T.V. Weather Forecasts*, Freakonomics.com, 12 April 2008
- **B. Hanson**, *Internet Weather Forecast Accuracy*, OmniNerd.com, 07 February 2007
- **A. Murphy**, *Climatology, Persistence, and Their Linear Combination as Standards of Reference in Skills Scores*, Weather and Forecasting 7: 692-698, 13 July 1992
- **J. Rosenberg**, *Accuweather long-range forecast accuracy questionable*, Josh-Rosenberg.com, 13 June 2013
- **H. Stern**, *The accuracy of weather forecasts for Melbourne, Australia*, Meteorol. Appl. 15: 65-71, 24 January 2008